

1 Postdoc research position in Natural Language Processing ERC project ENEP-CHINA

The Institute of Asian Studies (Irasia) at [Aix-Marseille University](#) invites applications for the following position: 1 Postdoc research position in the ERC project ENEP-CHINA.

This is a [European Research Council](#) (ERC) [Advanced Grant](#) project titled “Elites, networks, and power in modern urban China (1830-1949). Historical “big data” in modern Chinese history” headed by Prof. Christian Henriot. The funding period of the project is 1 September 2018 to 31 August 2023 (5 years).

The project ENEP-CHINA proposes a step-change in the study of modern China reliant upon scalable data-rich history. It will deliver precise historical information at an unprecedented scale from heretofore untapped sources - as well as reshaping the analysis of existing sources - to create a new dimension in the study of the transformation of elites in modern China. It will deploy an array of cutting-edge digital methods — including data mining, sampling, and analysis within an integrated virtual research environment. The key issue that the project will address is breaking through existing limits of access to historical information that is embedded in complex sources and its transformation into refined, re-usable and sustainable data for contemporary and future study of modern China. This project borrows substantially from the concepts and methods, especially quantitative methods and network analysis from sociology and political science. For the analysis of textual corpora, it draws on the procedures and applications developed in linguistics and literary studies to explore large corpora of unstructured texts. Finally, it implements the most advanced tools and methods designed in computing, especially data mining and machine learning.

The ENEP-CHINA project will be based at Aix-Marseille University on its Schuman campus (Aix-en-Provence) and will be implemented by an international interdisciplinary team. The project is also part of a large international network of researchers in Europe (Bristol, Lancaster, Leiden, Basel, Göttingen), China (Shanghai) and Taiwan (Academia Sinica).

Three research centers of [Aix-Marseille University](#) are involved in the ENEP-CHINA project:

[Institute of Asian Studies](#) (IrAsia)
[Laboratoire Parole et Langage](#) (LPL)
[Laboratoire d'Informatique et Systèmes](#) (LIS)

For further information about the principal investigator, visit [ankeqiang.org](#).

The position and what is offered

- The Postdoc research position is initially for 1 year with a possible extension of up to 6 months
- Starting date: preferably 1 September 2018
- Remuneration is based on Postdoc grade (100% working time) of Aix-Marseille University
- Salary scale : approximately 48,000 Euro per year before taxes and social security. The package includes including full social benefits, complementary insurance (mutuelle), and pre-paid meal tickets ((ticket restaurant).
- Joining an innovative research project funded by the prestigious ERC grant scheme
- Become a member of a large and fully interdisciplinary research team, including historians, linguists, computer and data scientists.
- Profiting from a vibrant internationally oriented research environment at Aix-Marseille University is available to attend international conferences, workshops, summer schools, and for research visits at institutions in the project countries

- No teaching obligations (but if desired, teaching opportunities can be offered in addition)

Your tasks

- Set-up a Natural Language Processing pipeline for Chinese and English historical texts (word segmentation, tagging, named-entities recognition, shallow semantic parsing)
- Apply Information Retrieval techniques on Chinese and English historical text to extract semantic relations
- Work in collaboration with engineers and data scientists of the team to a rich database from the “big data” unstructured textual data set
- Exploit reinforcement learning methods to take advantage of intermediate evaluation from data scientists, historians and linguists
- Publishing in peer-reviewed journals and conferences like ACL, COLING, EMNLP,...
- Supporting the ERC grant holder in managing tasks, coordinating the project work and supervising PhD students and student assistants

Qualifications and requirements

- Doctoral degree in Computer Science / Computational Linguistics (should be obtained before the start of the project)
- Publications in peer-reviewed journals on topics related to the field
- Experiences in one or more of the following fields: Machine Learning, Computational Linguistics, Information Retrieval
- Strong methodological skills in Programming, Scientific Computing
- Very good command of spoken and written English; French language skills are not required ; command of Chinese would definitely be an asset
- Ability to organize yourself efficiently and work collegially in an internationally oriented team

Employer : Aix-Marseille University

Aix-Marseille University (AMU) Aix-Marseille University offers training in all disciplines: Arts, Literature, Languages and Humanities; Law and Political Science; Economics and Management ; Health; Science and Technology. An intensive research site in collaboration with leading organizations, AMU is one of the French sites of excellence recognized by an international jury as part of Future Investments. The AMU Foundation helps develop research, training and employability of students, while creating a privileged link with the socio-economic world. Training, research, guidance, employability, interdisciplinarity and exploitation of knowledge are the pillars of this establishment, largest university in France and the Francophone world. For more information: <https://www.univ-amu.fr/en/university>

[Aix-Marseille University](https://www.univ-amu.fr/en/university) aims to increase the share of women in research and teaching and, therefore, explicitly encourages applications from women. As an international research institution, Aix-Marseille University particularly welcomes applications from abroad.

Your application must include:

- CV – including degrees and other completed courses, work experience and a list of degree projects/theses, language skills
- A letter (max. 1000 words, excl. references) describing your research interests, skills and motivation to apply for the position; describe any digital method(s) you have used
- A copy of relevant degree(s)
- Two work samples where you used digital methods (e.g. journal articles, thesis or book chapters, etc.)

- Two reference letters with contact details [**The letters of reference must be sent directly by the referees to Prof. Christian Henriot at contact@numerica-sinica.eu.**]

Please send your application (motivation letter, CV, copy of your PhD certificate/transcripts and work samples in one PDF document to contact@numerica-sinica.eu. Use the subject heading “ENEP-CHINA NLP Postdoc Research Position”.

Deadline: 15 July 2018, 12:00 noon CET.

If you have any further enquiries about the positions, please do not hesitate to contact Prof. Christian Henriot at christian.r.henriot@gmail.com